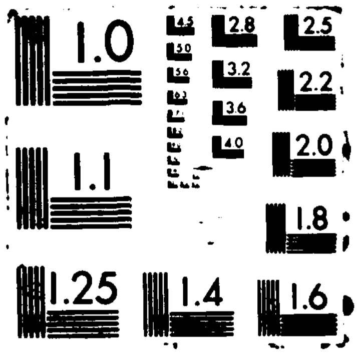
AD-A181 725 WORKSHOP ON COMPLEX SOUND PROCESSING HELD IN SARASOTA FLORIDA ON 26-28 APRIL 1986(U) LOYOLA UNIV OF CHICAGO IL W YOST 28 APR 87 AFOSR-TR-87-0710 AFOSR-85-0351 1/1

UNCLASSIFIED F/G 6/4 NL


1.0
2.8
2.5
3.2
2.2
3.6
4.0
2.0
1.1
1.8
1.25
1.4
1.6

UNCLAS
SECURITY CLA

DTIC FILE COPY

(2)

AD-A181 725

REPORT DOCUMENTATION PAGE

DTIC ELECTE JUN 1 6 1987 S D D

1a. REPORT SECURITY CLASSIFICATION: UNCLAS	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited.
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 87-0710
6a. NAME OF PERFORMING ORGANIZATION: LOYOLA UNIVERSITY OF CHICAGO / 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: Air Force Office of Scientific Research/NL
6c. ADDRESS (City, State, and ZIP Code): 820 N. MICHIGAN AVE CHICAGO IL 60611	7b. ADDRESS (City, State, and ZIP Code): Building 410 Bolling AFB, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR / 8b. OFFICE SYMBOL (If applicable): NL	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: AFOSR-85-0351
8c. ADDRESS (City, State, and ZIP Code): Building 410 Bolling AFB, DC 20332-6448	10. SOURCE OF FUNDING NUMBERS: PROGRAM ELEMENT NO. 61102F; PROJECT NO. 2313; TASK NO. A6; WORK UNIT ACCESSION NO.

11. TITLE (Include Security Classification)
(U) WORKSHOP ON AUDITORY PERCEPTION OF COMPLEX STIMULI

12. PERSONAL AUTHOR(S) Dr. William Yost

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
FINAL	FROM 30 SEP 85 TO 29 SEP 86	87 APR 28	5

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	HUMAN AUDITION/PSYCHOPHYSICS
05	09		

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

A brief synopsis of a workshop held in Sarasota FL, 26-28 ~~March~~ April 1986. Complete manuscripts are available in the book "Auditory Processing of Complex Sounds", Wm Yost + CHARLES WATSON (Eds), Lawrence ERLBAUM: Hillsdale NJ (1987); ISBN 0-89859-981-4.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: ☒ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL: JOHN F. TANGNEY	22b. TELEPHONE (Include Area Code): (202) 767-5021 / 22c. OFFICE SYMBOL: NL

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
UNCLASSIFIED

Final Technical Report
on the Workshop on Complex Sound Processing
supported by
the Air Force Office of Scientific Research, Life Sciences
(AFOSR 85-0351)

The workshop was supported by the Air Force Office of Scientific Research (AFOSR), Life Sciences, and was chaired by the editors of this book. A series of recent events led to the workshop and publication of this book. In 1982, Dr. John Tangney of AFOSR approached the Committee on Hearing, Bioacoustics and Biomechanics (CHABA) of the National Academy of Sciences to survey recent developments and trends in the study of the auditory system. The result of the request from AFOSR was a 1983 Symposium on Basic Research in Hearing organized by CHABA and sponsored by the AFOSR (Dolan and Yost, J. Acoust. Soc. Am. 78, No.1 Part 2, 1985). After reviewing the proceedings of the CHABA Symposium and considering its program goals, AFOSR began, in 1985, a program of support for research on complex auditory perception. The support by the AFOSR, the discussions at the CHABA Symposium, and the increased volume of research on the topic of auditory processing of complex sounds stimulated us to organize a meeting on this topic. With the support of the AFOSR the Sarasota Workshop on Auditory Processing of Complex Sounds was held in April, 1986.

Thirty scientists presented papers at the workshop and another fifteen scientists attended as observers. Three days of papers and discussion took place. We did not organize the workshop with the intent of publishing a book. The topics were chosen from the many excellent submitted papers in order to sample as diverse a cross-section of research as possible and yet provide continuity to the three-day meeting. The quality and quantity of abstracts submitted for inclusion in the workshop and the enthusiastic and insightful discussions at the meeting convinced us and the participants that a timely publication devoted to these topics would be a useful contribution. Therefore, following the workshop the authors prepared chapters in camera-ready form in order to produce a book in a short period of time. The chapters are not just transcriptions of the presentations given at the workshop, but they are written as brief papers on the topic of the author's interest. Authors were encouraged to provide a brief background to their work and to make sure the germinal references on their topic were included in their bibliography. The book, Auditory Processing of Complex Sounds, was published by Erlbaum Press in early 1987 and is now available.

Assistance for the project came from many sources. The Workshop and the book would not be possible without the foresight and dedicated support of John Tangney as a Program Director in the Life Sciences Division of AFOSR. The staff at the Sarasota Sheraton Hotel provided a pleasant environment in which to meet. Lawrence Erlbaum Press has been very helpful in assisting us in getting the book out quickly. The staff of the Parmly Hearing Institute at Loyola University, especially Marilyn Larson, Beth Langer, Ned Avejic, and Scott Stubenvoll have been invaluable, as has the staff of the Department of Speech and Hearing Sciences at Indiana University, especially Janet Farmer.

Below is a summary of the major topics discussed at the workshop and contained in the book.

This workshop brought together investigators with a remarkable diversity of approaches to the general problem of how humans (and nonhumans) process (or "hear," or "perceive") complex sounds. The only common denominator at the onset was that each had responded to an announcement (mailed or published in a journal), asking for contributed papers for a "workshop on complex sound perception." Surprisingly, this yielded a range of topics, research paradigms, and theoretical perspectives with some well-defined themes.

We anticipated that "complexity" would mean different things to different people, but the range of meanings that can be inferred from these twenty-eight papers is actually relatively small. In general, "simple sounds" are considered to be the individual pure tones or noise bursts that have served as the stimuli in most studies of the auditory system since Helmholtz. "Complex stimuli" mean those that vary systemically in either their spectrum, or in time, or both. While most of the contributors created complex stimuli to test particular hypotheses about auditory processing, a few dealt with natural or environmental sounds, speech, birdsongs, or music.

Many of the authors avoided the need to discuss physical criteria for stimulus "complexity," and instead opted for distinctions based on mechanisms of processing. "Simple processing" in the spectral domain was equated by most authors with a critical band (CB) model, and in the temporal domain with the time constant of a simple temporal integrator. "Complex processing" was shown to require a considerable variety of mechanisms beyond these more traditional workhorses of auditory theory, including spectral-shape and temporal-pattern detectors, and even more elaborate mechanisms (hardware, software, or both) whose operation in many cases requires knowledge of the sources of complex sounds.

In general, the contributions can be divided into: (1) spectral processing, (2) temporal processing, (3) pitch, (4) speech, (5) physiological processing, and (6) perceptual organization; including "object" or event perception and central mechanisms. These a posteriori categories cannot, of course, capture the scope of numerous papers that treated more than one of these topics as, for example, several that dealt with stimuli varying both in spectrum and in time. The papers have been grouped into these six categories, but the reader is warned not to expect discussions of spectral processing to be confined to papers in the section bearing that name, and-so-on.

The chapters that deal with spectral aspects of complex processing generally agree, as observed above, that considerably more elaborate frequency analysis can be demonstrated in psychoacoustic experiments than is predictable from a "bare-boned" critical-band filter bank. It should be stressed that none of these "failures of critical band theory" in fact provide evidence against the CB as an initial stage in frequency analysis. Several lines of investigation, however, demonstrate that when it is to the advantage of the listener to do so, he or she can simultaneously process energy arriving in several critical bands. That ability is demonstrated in two types of experiments. In one, a broad-band spectral array itself is treated as the meaningful event (a "signal"), rather than just one part of the spectrum (that associated with the output of a single auditory filter). Studies of "profile analysis" or spectral shape discrimination and its derivatives are examples of this approach. In the other, it is shown that temporal correlations among the noise levels across critical bands can reduce the masking efficiency of a critical-band masker (co-modulation release from masking or CMR). In both cases, mechanisms are implied which are simultaneously sensitive to the relative levels in each of a number of adjacent auditory channels. Common sense would have predicted at least one of these findings; vowel identification obviously requires recognition of spectral shape. Some of the chapters discuss the nature of the physiological code that might subserve spectral pattern processing. The consensus seems to be that rate codes <u>and</u> temporal codes are both used by the central nervous system to process complex spectral patterns. These lines of research (both psychophysical and physiological) promise to establish the limits within which such spectral shape- or profile-based recognition can operate.

Many sounds of everyday life may be described as temporal sequence of stimuli. If very similar (highly correlated) sounds occur in close temporal proximity, then under many circumstances, the auditory system is most sensitive to the first arriving information rather than to the pattern of the events. Studies of the precedence effect have provided insights into the mechanisms that govern the influence of the first acoustic wavefront. When the sequence of sounds is made up of different or uncorrelated acoustic events, the temporal pattern may lead to a variety of perceptions. Often times one part of a temporal pattern may be "heard out" from the background of the rest of the sound. In many contexts the last acoustic events are the most salient. The analogy to the foreground/background concepts of stream segregation (as derived from Gestalt Theory) is one theoretical approach to describe the dominance or saliency of certain aspects of a complex temporal pattern. Several computational schemes also provide insights into how to model discrimination among different sequences of sound. A variety of lines of research show the major role played by temporal modulation in our perception of complex sounds. The abundance of useful information available in the temporal code of the auditory nerve provides a physiological argument favoring temporal modulation as a variable around which many perceptions of complex sounds appear to be organized.

There are only so many words that can be used to describe a sound. One of the most common words is "pitch." Although there is some disagreement about the precise definition of pitch, a variety of complex sounds are capable of producing sensations listeners refer to has having pitch. Many authors consider pitch to be a major organizing feature for our perceptions of complex sounds. Models based only on auditory neural tuning or only on neural temporal periodicity, have failed to provide adequate descriptions of the pitch evoked by many complex sounds. Thus, the debate concerning whether complex pitch is spectrally or temporally based continues. Much of the research in this book suggests that the extraction of pitch from complex stimuli is not an "either-or" question. In both spectral shape processing and pitch processing, neural tuning _and_ temporal coding must be considered. In addition, although the auditory nerve contains a wealth of temporal and spectral information, central mechanisms might be required to fully process the peripheral neural code in a manner adequate to account for complex pitch perception.

If a complex sound contains short term spectral changes then these might give rise to pitches which listeners could use in processing these sounds. The work on stream segregation, spectral shape discrimination, and tonal pattern recognition emphasizes the need to consider carefully possible long-term and short-term spectral cues that may be used to detect, discriminate, or identify many complex sounds.

A lot of the work generally concluded, not only that the peripheral mechanisms of auditory tuning and simple temporal integration are inadequate to explain the hearing of complex sounds, but also that some fairly elaborate central processing must be involved. A few papers explicitly deal with selective attention, short-term memory capacity, and other such cognitive constructs. It is clear that the "passive" auditory system is in fact very dynamic and can effectively be "programmed" to look like quite a variety of acoustic information processing devices. If we are to cope with such practical issues as auditory code learning (speech or non-speech), it is essential that we learn some of the primary limitations within which the central processor functions. How long can a sound be, if it is to be accurately recalled, or recognized later? How much of a complex sound must be processed "categorically," if any? Within what parameters must selective auditory attention function? Are there two auditory modes, one for speech and one for non-speech? Or, do we process very familiar sounds (e.g., speech in our native tongue) differently from novel sounds? Several papers made efforts to deal with these issues, but it is clear that a great deal remains to be done before we will understand the actual auditory processing that occurs at a cocktail party.

One fascinating line of thought carries on from the tradition of Gestalt Psychology. Certain organizing principles seem to be used when we hear a novel complex sound. Sometimes a portion of a total waveform "stands out", i.e., seems to be closer. <u>That</u> is an instance of auditory Gestalt perception, and many such studies must be collected to determine the organizing principles with which listeners deal with most novel sounds. Those principles will certainly include frequency similarity as one of the most potent determinants of a "figure." It appears that musicians may be ahead of basic scientists in this area. Many of these concepts appear to be applicable whether we use speech and human communication, complex non-speech sounds, music, or an animal model, such as songbirds, as our tool for understanding auditory processing.

Materials and Publications relevant to the workshop have been sent over the past year to the AFOSR Program Officer, Dr. John Tangney.

END

7-87

DTIC